

RESEARCH

COMPARISON OF GRIDDED AND MEASURED RAINFALL DATA FOR BASIN-SCALE HYDROLOGICAL STUDIES

Enrique Muñoz^{1*}, César Álvarez², Max Billib³, José Luis Arumí⁴, and Diego Rivera⁴

Global gridded climatological (GGC) datasets, including precipitation and temperature, are becoming more and more precise, accessible, and common, but the utility of these datasets and their limits for hydrological research are still not well determined. In this paper, we compare the performance of two hydrological models that are identical in structure but built with two different inputs: rainfall from rain gauge stations and from a GGC dataset. The objective is to evaluate the utility of gridded datasets in water resource availability studies mainly for hydroelectric and agricultural purposes. The Andean basin of the Laja River, located in south-central Chile, was chosen for this study. It was based on an 18-yr simulation, and it was concluded that i) with gridded climatological datasets in a monthly water balance model, it is possible to reproduce the behavior of an Andean basin with good goodness-of-fit, but with worse results than when using inputs from rain gauges; ii) the amount of rainfall in gridded datasets in the Andean area of the Laja basin is underestimated and damped, an effect which is transferred to the simulated flows; and iii) regarding the main activities in the Laja basin, global gridded datasets are useful for hydrological studies with agricultural purposes prior to a treatment that considers the orographic effect. On the other hand, these datasets are useless for hydroelectric purposes due to the large underestimation of peak flows obtained during the rainy season.

Key words: Flow simulation, monthly water balance, surface water hydrology, water resources.

Irrigated agriculture and hydroelectric power are two of the main economic activities in Chile's Central Valley, and both depend on the availability of water from Andean watersheds. In recent years, Chilean water users have been expressing their concern about the impact that climate variability and land-use changes might produce on water availability. On the other hand, despite the fact that the south-central area of Chile is not yet an area of water scarcity (Rijsberman, 2005; Oki and Kanae, 2006), population growth and development are increasing the demand for water resources and their more efficient administration and use.

The lack of streamflow data in many watersheds has made it necessary to use monthly hydrological simulation models to evaluate irrigation water availability or the impacts of large irrigation projects on hydrology. Water balance calculations are currently one of the most important ways to perform both climate and hydrological

research since they can provide crucial information about the amount of water circulating in the hydrological cycle, the amount of renewable water available for ecosystems and human use, and the estimation of impacts on water resources due to climate change and variability.

In a simple water balance, water storage in a basin can be expressed as $R = P - E - dS/dt$ where P is precipitation at time dt [$L T^{-1}$], E is evapotranspiration at time dt [$L T^{-1}$], dS/dt is the change in basin water storage [$L T^{-1}$], and R is the excess rainfall, which conceptually is the loss from the basin from both surface and subsurface storage produced at time dt [$L T^{-1}$]. In the above equation, P is the most important variable in the balance because it is normally the only water input in the basin, and the other balance variables depend directly or indirectly on P , hence, the water balance will be conditioned by the amount of rainfall.

Precipitation is the main input in the hydrological system. However, rain gauge climatological data, such as temperature among others, in mountainous regions as in the Andes, are scarce or even nonexistent, and this is a problem that hinders modeling. For example, Stehr *et al.* (2008) used the Soil and Water Assessment Tool model to estimate the impact of climate change on the Bio Bio River Basin, but the study was limited by the lack of soil data and precipitation in the Andes. Vicuña *et al.* (2011) had a similar problem with a lack of rainfall data above 3000 m.a.s.l. in a study on Andean basins. On the other hand, due to the increase in computational capabilities and the related easier access to measured meteorological

¹Universidad Católica de la Santísima Concepción, Facultad de Ingeniería, Alonso de Ribera 2850, Concepción, Chile.
*Corresponding author (emunozo@ucsc.cl).

²Universidad de Cantabria, Environmental Hydraulics Institute "IH Cantabria", Avda. de los Castros s/n. Santander, España.

³Gottfried Wilhelm Leibniz University of Hannover, Institute of Water Resources Management, Appelstr. 9a, D-30167, Hannover, Germany.

⁴Universidad de Concepción, Facultad de Ingeniería Agrícola, Av. Vicente Méndez 595, Chillán, Chile.

Received: 2 March 2011.

Accepted: 22 June 2011.

information all over the world, it is far more frequent to use long-term and world-scale rainfall and temperature data where both are structured/presented by grids. Reliable methods for climatological data interpolation/extrapolation in ungauged regions would address a critical weakness in hydrological modeling of unmeasured basins, thus providing a potential source of input data to force basin-scale hydrological models.

This study aims to examine the deficiencies and evaluate the utility and reliability of interpolated global scale data as an option to supplying data that is lacking for hydrological studies, in particular on studies of basin-scale water resource availability for agricultural, hydroelectric, environmental, and/or human purposes. Deficiencies, utility, and reliability were evaluated by comparing two hydrological models, which used gridded and measured datasets as model inputs.

MATERIALS AND METHODS

Case study

The Laja River has a multiple-use snow-rain regime and a complex interaction among natural, economic, and social components that control water resources (Mardones and Vargas, 2005). In this type of mountainous watershed with abundant water and forest resources, there is a special ecological and economic interest because it is essential in territorial and national energy plans.

The Laja basin (Figure 1) covers an area of 4668 km²; it is bounded to the east by the Andean Mountains, to the north by the Antuco-Sierra Velluda volcanic complex, to the south by the Chillán Volcano, and to the west by the Cordillera de la Costa (Coastal Range) where it ends in the Bío Bío River. Its elevation ranges from 50 to 3534 m.a.s.l. The Andean zone of the river (600 to 2000 m.a.s.l.) is composed of sedimentary and plutonic rocks on which there are volcanic and volcanoclastic deposits of the Antuco-Sierra Velluda volcanic complex (Thiele *et al.*, 1998). The central depression consists of the large Laja River black sand cone alluvial deposit (100 to 300 m.a.s.l.), which is above a lahar formation. These deposits end at the Cordillera de la Costa (Coastal Range), which is composed of granite (this limit represents the northwestern edge of the Claro River). The combination of sand and lahar deposits yields a permeable soil layer over an impermeable base that promotes the occurrence

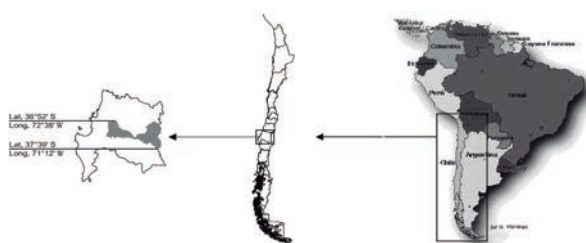


Figure 1. Location of the Laja River basin.

and maintenance of groundwater reserves in the central depression. Moreover, the interaction of alluvial deposits with the Coastal Range affects the shape, orientation, and hydrology of rivers (due to the impervious nature of the Coastal Range with its high slopes that are an impermeable barrier for groundwater flow and a runoff accelerator for direct runoff), an effect observed in the Claro River, a tributary of the Laja, that has a north-south orientation and high, constant year-round flows.

This basin has been historically affected by anthropogenic alterations that have sought the use of water resources for socio-economic purposes to the point of over-demand (Muñoz, 2010). Furthermore, due to the spatial variability of the hydrological responses in the Laja Basin, the relative importance acquired by its dominant processes and the anthropogenic processes affecting it, the Laja River basin is viewed as a good candidate to evaluate world-scale gridded climatological datasets for hydrological purposes. The Laja River Basin exhibits a high spatial variability in its hydrological behavior, is highly influenced by the snow component, high slopes, and mostly impervious soil layers on the upper third, as well as by the pluvial component, medium to small slopes, and granular soil layer that favors groundwater flows on the central and lower third. Consequently, the river and its hydrological behavior across the Laja basin must be discretized and characterized through sub-basins whose outflow is controlled by a fluviometric station.

Description of water balance model

The main factors to consider in selecting a model are the purpose for which the model is required and resource availability (in terms of time, data, computing facilities, etc.). In the case of a scarce-data basin, a simple conceptual water balance model is the best option. Given that the conceptual pluvial models do not differ a lot one from the other for conceptualization and capabilities (Jiang *et al.*, 2007), the Ferrer *et al.* (1973) model (hereafter, the BFA model) was chosen as the basis for this study.

The BFA model has been widely used in Chile, mainly in applied hydrology for agricultural and engineering purposes, and has been broadly accepted by the local hydrological community. This model represents the basin as a double storage system (Figure 2, pluvial model), a sub-superficial (unsaturated) and a subterranean (saturated) system. The sub-superficial layer is supplied by rainfall infiltration (INF) although the subterranean layer is supplied through direct percolation from the basin surface and/or from the upper storage system. Emptying the storage systems produces the baseflow, which added to surface runoff, constitutes basin outflow. The BFA model is a pluvial model where the inputs are monthly precipitation (MP) and potential evapotranspiration (PET). The only output is Total Runoff (TOTR) at the basin outlet, including both Direct Runoff (DR) and Baseflow (BF). The model has six parameters, plus two

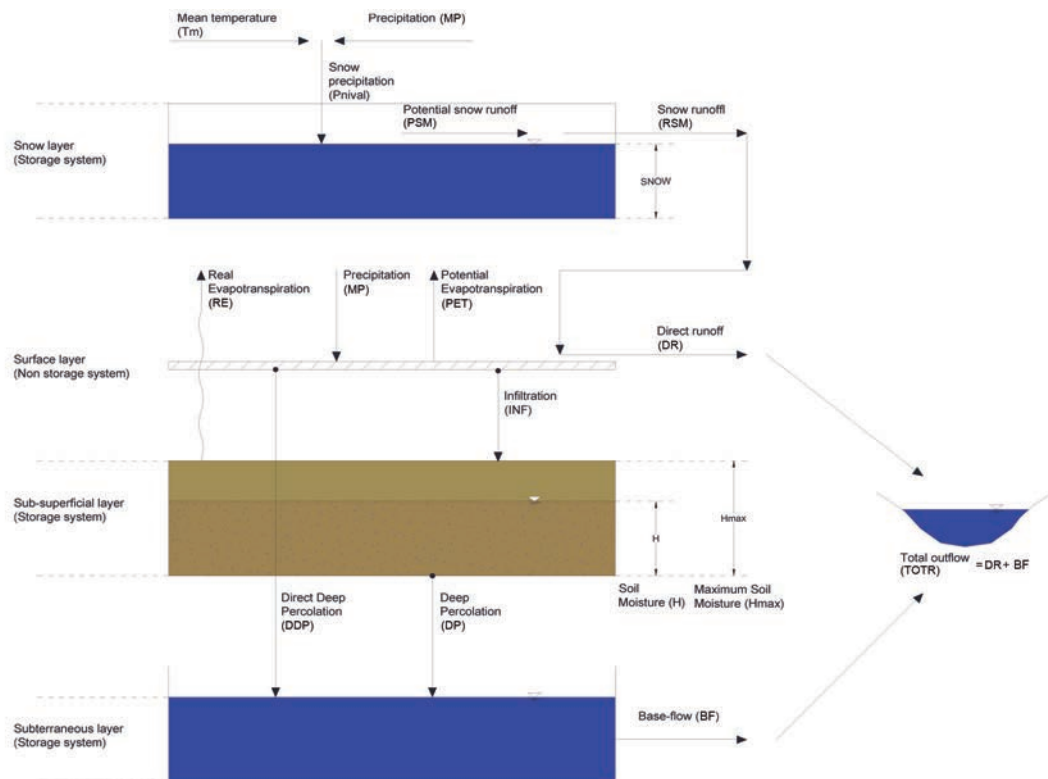


Figure 2. Conceptual diagram of the pluvial and snowmelt model.

others that allow adjusting the inputs described in Table 1.

By considering the limitations of the BFA model, and the basin needs under study, the model was modified by coupling a snow-melt model to it and a flow alteration model (that allows including alterations, such as irrigation and transfer channels, and hydroelectric operations through a flow data series).

The snow-melt model is coupled as an external module where the output (RSM) is runoff from the snowmelt and incorporated into the pluvial model by adding a portion of RSM to DR, and the rest to INF. This portion is defined through the calibration F parameter (Table 2).

Table 1. Description of BFA model parameters and adjustment factors and their influence on other variables.

| Parameter | Description | Influence |
|----------------|---|-------------------|
| C_{max} | Maximum runoff coefficient when the sub-surface layer is saturated | DR |
| P_{lim} , mm | Rainfall limit over which there is DDP | DDP |
| D | Rainfall percentage on P_{lim} transformed into DDP | DDP |
| H_{max} , mm | Maximum capacity of the soil layer to retain water | C_{max} and RE |
| PORC | Fraction of H_{max} that defines soil water content restricting the evaporation processes | H_{crit} and RE |
| C_k | Subterranean runoff coefficient | BF |
| A | Precipitation data adjustment | MP |
| B | Potential evapotranspiration data adjustment | PET and RE |

BFA model: Ferrer *et al.* (1973) model; DR: Direct runoff; DDP: Direct deep percolation; RE: Real evapotranspiration; BF: Baseflow; MP: Monthly precipitation; PET: Potential evapotranspiration.

Conceptually, this model consists of a virtual basin located upstream from the pluvial basin, and it is represented as a simple storage system with only a snow regime. The inputs are MP and mean monthly temperature (T_m) from which snowfall (P_{nival}) is calculated. P_{nival} recharges the snow layer, and based on T_m , potential snowmelt (PSM) is calculated. Then, depending on stored snow, real snowmelt (RSM) is calculated, i.e., if there is less snow than the potential snowmelt, RSM will be stored snow, otherwise actual and potential snowmelt are the same. Later, the RSM output is incorporated into the pluvial model in DR and INF. Figure 2 shows the conceptual diagram of this model integrated with the BFA pluvial model, and Table 2 shows a description of parameters related to the snowmelt module.

Table 2. Snowmelt model parameter description.

| Parameter | Description | Influence |
|------------------------|---|-------------|
| FgT | Factor that modifies the thermic gradient (should be 1 if thermic gradient is assessed) | P_{nival} |
| M, mm °C ⁻¹ | Fraction of snowmelt over a base temperature (T_b) where melting starts | PSM, RSM |
| T_b , °C | Value of temperature over which melting starts (usually 0 °C) | PSM, RSM |
| MM | Minimum rate of melting when $T_m < T_b$ | PSM, RSM |
| F | Fraction of the real snowmelt which goes to DR | DR |

P_{nival} : Snowfall; PSM: Potential snowmelt; RSM: Real snowmelt; DR: Direct runoff.

On the other hand, the flow alteration model allows incorporating basin alterations that modify the normal streamflow regime. Furthermore, it allows discretizing the basin into sub-basins where the outflow of one sub-basin is the inflow of another located downstream. With that change, the model is capable of simulating a basin through various sub-basins with different geomorphological and hydrological characteristics, as well as including alterations in the river basin. This model is as follows:

$$Q_{\text{out}}(t) = Q_{\text{sub-basin}}(t) + Q_{\text{contributions}}(t) - Q_{\text{extractions}}(t) \quad [1]$$

where Q_{out} is the sub-basin outflow, $Q_{\text{sub-basin}}$ is the flow produced by the sub-basin (equivalent to the TOTR in the BFA model), and $Q_{\text{contributions}}$ and $Q_{\text{extractions}}$ are the inflows/outflows to/from the sub-basin due to alterations. See Muñoz (2010) for a detailed description of snowmelt and flow alteration models.

Evaluation and selection of global gridded datasets

The following global monthly gridded climatological datasets were reviewed:

a) Climate Research Unit (CRU) (New *et al.*, 1999; 2000; CRU, 2000): This database includes various climate datasets, such as precipitation, temperature, pressure, and circulation indices. This dataset gathered station data from formal and informal sources, which was later interpolated through the thin-plate spline method. Dataset resolution was 5×5 and 2.5×3.75 (latitude \times longitude in geographical coordinates) for the 1900-1998 period.

b) Global Precipitation Climate Center (GPCC) (GPCC, 2010): This product is based on quality-controlled data from 7000 to 8000 stations, 6500 synoptic ocean prediction and infrared radiometers to measure atmospheric and ground surface stations. Precipitation datasets are available from 1986 to the present in a 1×1 and 2.5×2.5 resolution.

c) Global Precipitation Climatology Project Version 2.1 (GPCP) (GPCP, 2010): This project developed precipitation datasets from remotely sensed data with microwave and infrared sensors from geostationary and polar-orbiting satellites along with ground observations. Precipitation datasets are available from 1979 to 2009 with a 2.5×2.5 resolution.

d) Department of Energy of the National Center of Environmental Prediction – Reanalysis II (NCEP/DOE R-2) (Kanamitsu *et al.*, 2002; NCEP, 2004): This dataset was obtained from observed data such as from meteorological stations and ships, among other sources. The information is available from 1979 to 2004 with a 2.5×2.5 resolution.

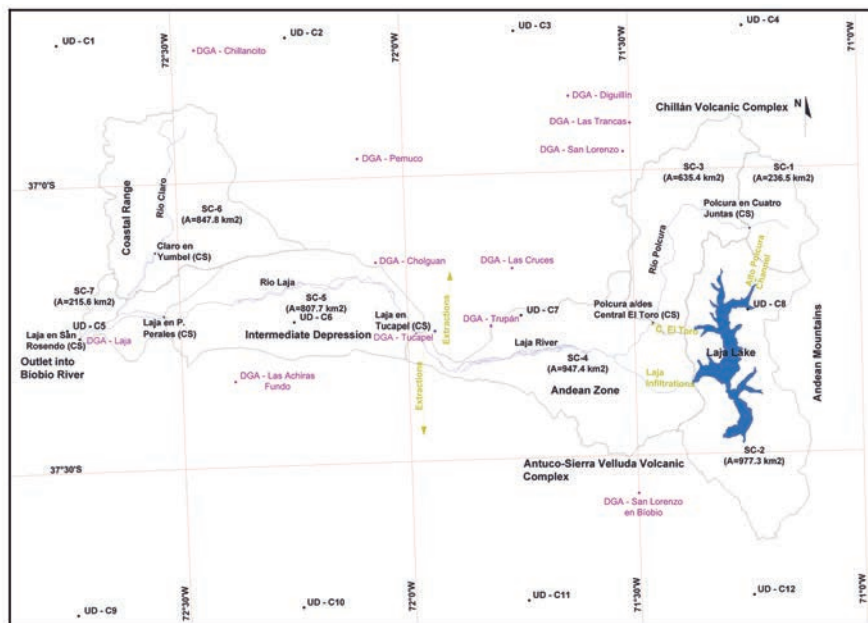
e) University of Delaware (UD) (Matsuura and Willmott, 2008): This dataset is available for the period 1960 to

2008 (for South America) with a 0.5×0.5 resolution. Interpolation is calculated from 5315 stations from the internal “South American Precipitation: Station Records Archive (version 1.02)” file. Interpolation is conceptually based on the idea that large-scale spatial variation of climate variables is more important than the interannual variation; this leads to the assumption that it of greater importance to incorporate all available stations in the interpolation than it is to maintain rigorous time series consistency.

Pitcher *et al.* (1983) concluded that rain gauge-averaged datasets are the most reliable large-scale data compared with remotely sensed and general circulation model estimations. Willmott and Legates (1991) later mentioned that long-term, large-scale precipitation means from historical rain gauge data have been accepted as unbiased by many of the climatological and hydrological communities. More recently, in a study about precipitation uncertainties and the related impact on runoff estimates, Fekete *et al.* (2004) reviewed different gridded climate datasets and concluded that: i) the GPCC and GPCP datasets appear to be very different in terms of annual mean precipitation, are similar in mid-latitudes, and diverge mostly in high latitudes; ii) the CRU and UD datasets are very similar in all latitudes; iii) the NCEP/DOE R-2 dataset greatly differs from other datasets presented here in terms of the amount of mean annual precipitation and mean absolute deviation; and iv) the most similar data analyzed were GPCC, GPCP, CRU, and UD; NCEP/DOE R-2 were excluded from the study due to their extreme anomalies compared to the other four datasets. Thus, based on the fact that both are estimated from rain gauge stations, we can assume that the most reliable datasets are those produced by CRU and UD. Given that the UD dataset is more precise (better resolution) and its data is more recent, it was chosen for the water balance calculations.

Model inputs

The pluvial model requires monthly precipitation and potential evapotranspiration as inputs; in the case that the basin is modeled under a snow-rain regime, it also requires the sub-basin monthly mean temperature. There are only rainfall data measured from the Dirección General de Aguas (DGA) stations Chilean water resources administration agency), as well as rainfall and temperature data from the UD datasets. Two input datasets were built based on this, one with rainfall from rain gauges and the other with global scale gridded precipitation (hereafter DGA and UD inputs, respectively). In both cases, potential evapotranspiration was estimated with the Thornthwaite method (Thornthwaite, 1948) based on the UD temperature datasets, which were also input for the snow module. Spatial distribution of the meteorological variables by sub-basin was achieved through Thiessen polygons. See Figure 3 for the location of the DGA stations and UD center grids.



DGA: Indicates rain gauge stations; UD: Indicates center grid locations.

Figure 3. Conceptual model of the Laja basin and its discretization. DGA/UD stations are indicated by light/dark dots, rivers and sub-basin limits are shown with dark lines and external alterations in light-colored lines.

Model calibration, validation, and simulations

The only period of flow records in common across the basin (including channels) is between August 2003 and June 2006; for that reason, calibration and validation were carried out for the periods Aug. 2003 to Dec. 2004 and Jan. 2005 to June 2006, respectively. Because SC-6 is independent from the rest of the system, calibration and validation for this sub-basin were performed between Jan. 1986 to Dec. 1989 and Jan. 1991 to Dec. 1995, respectively.

Calibration was performed manually (trial and error) by Pearson's correlation coefficient (R^2) as the objective function, single set of parameters was selected that maximizes both models, and also has a conceptual physical meaning (based on the sub-basin characteristics and their spatial variability). The only parameters defined as open to get different values were parameters A and B (to adjust rainfall and evapotranspiration data, respectively).

Subsequently, a simulation of the period between Jan. 1990 and Dec. 2008 was carried out based on the parameters obtained by calibration. Moreover, to support the analysis and discussion about the simulation period, four model performance assessments were calculated: i) Relative root mean squared error (RRMSE), ii) correlation coefficient (R^2), iii) coefficient of efficiency (NSE) (Nash and Sutcliffe, 1970), and iv) percent bias (PBIAS).

RESULTS AND DISCUSSION

Model application

Sub-basin model. The Laja basin was conceptualized as

a seven sub-basin semi-distributed model (Figure 3) from SC-1 (upstream) to SC-7 (downstream). It is important to mention that Laja Lake is a natural lake, but its outflow is artificially controlled by hydroelectric generation (with a mean operation flow of $34 \text{ m}^3 \text{ s}^{-1}$) released into the Polcura River through the El Toro discharge. There are also exfiltrations from Laja Lake concentrated in Los Ojos del Laja, starting point of the Laja River with a mean flow of $27 \text{ m}^3 \text{ s}^{-1}$. A brief description of each sub-basin follows:

a) SC-1, Polcura Alto: This is a snow-rain sub-basin with two outflows, one through the Canal Alto Polcura (a transfer channel), which goes to Laja Lake, and the other through the Polcura River up to the Polcura in Cuatro Juntas fluvimetric station. It is an Andean basin with an elevation between 1350 and 2910 m.a.s.l. and where a low-permeability soil layer prevails.

b) SC-2, Lago Laja: Laja Lake is located on the Andes Mountains slope on a low-permeability depression of volcanic soil. The basin has two outflows, the El Toro hydroelectric station and Ojos del Laja (exfiltrations), and both are in SC-4.

c) SC-3, Polcura: This basin is located in the mountainous area of the Chillán Volcano between 800 and 3089 m.a.s.l.; it has a low-permeability soil layer, and high slopes ($\geq 1.6\%$). It has input from SC-1 and discharge in SC-4, and it is controlled by the 'Polcura Antes de Descarga Central El Toro' fluvimetric station.

d) SC-4, Laja Alto: This is the most complex sub-basin in the whole system. It receives three inputs and has eight extractions (Zañartu, Collao, Mirrihue, El Litre, Bulnes, Ortiz, Laja-Diguillín, and Laja Sur channels). It is a snow-rain Andean sub-basin located between 289 and 3534 m.a.s.l. with medium slopes and mainly composed of volcanic rocks and granular deposits. The outlet is controlled by the Laja en Tucapel station, which also divides the upper and lower parts of the basin.

e) SC-5, Laja Medio: This sub-basin can be modeled as a pluvial basin (there is no important snow influence) with inflow from SC-4 and outflow into SC-7. The control station is the Laja en Puente Perales station, which has a continuous series during the simulation period. It is located in the central depression where sandy deposits on a lahar formation prevail, benefiting aquifer and groundwater flows.

f) SC-6, Río Claro: This sub-basin is independent from the rest of the system. It has no superficial inflows; it is controlled by the Río Claro Camino a Yumbel fluviometric station and discharges into SC-7. Regarding the geology and geomorphology, it is a particular basin because the area from the slope of the Coastal Range to the Claro River is mountainous, forested, and composed of granite (low permeability). Meanwhile, the area located to the east of the Claro River has small slopes (0.3 to 0.8%) with higher permeability, low forest density, and mostly used for agricultural purposes.

g) SC-7, Río Laja en Desembocadura: This sub-basin has two inflows (from SC-5 and SC-7) and has no fluviometric control. Therefore, model performance is not evaluated at this point.

Calibration, validation, and simulation. The resultant correlation coefficients obtained from calibration are shown in Table 3. SC-2 and SC-7 sub-basins have no values because SC-2 is Laja Lake, which is not modeled in this study (outflows are based on the lake level and on hydroelectric demand), and SC-7 has no flow records for comparison.

Table 3. Correlation coefficient for model calibration (Aug. 2003 to Dec. 2004) and validation (Jan. 2005 to June 2006). Except for SC-6, calibration/validation was for the period Jan. 1986 to Dec. 1989/Jan. 1991 to Dec. 1995.

| R ² | Calibration | | Validation | |
|----------------|----------------------|---------------------|----------------------|---------------------|
| | DGA _{input} | UD _{input} | DGA _{input} | UD _{input} |
| SC-1 | 0.68 | 0.65 | 0.75 | 0.74 |
| SC-3 | 0.91 | 0.86 | 0.95 | 0.87 |
| SC-4 | 0.87 | 0.77 | 0.84 | 0.75 |
| SC-5 | 0.96 | 0.95 | 0.94 | 0.92 |
| SC-6 | 0.94 | 0.92 | 0.95 | 0.94 |

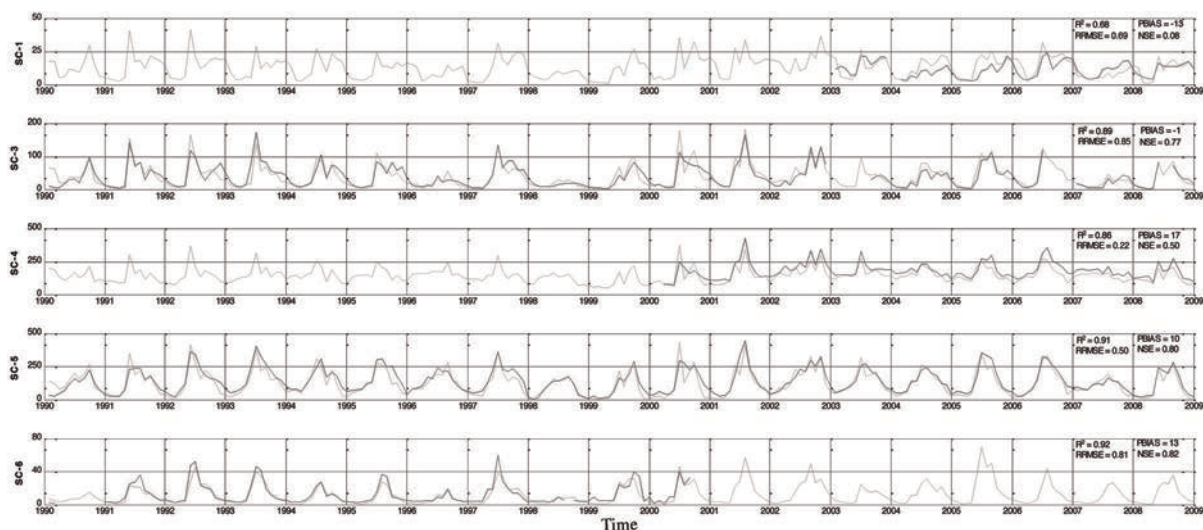
R²: Pearson's correlation coefficient; DGA_{input}: An input model made up of rainfall data and gridded temperature data measurements; UD_{input}: An input model made up of gridded rainfall and temperature data.

Figures 4 and 5 show simulation results and compares simulated and observed flows. Moreover, figures show the four model performance assessments calculated for the simulation period.

Both models (with DGA and UD inputs) accurately reproduce the snow-rain regime, which is mainly observed in sub-basins SC-1 and SC-3 (Figures 4 and 5). It is possible to identify two peaks, one during the winter season and the other during early summer. It is also observed that the snowmelt influence decreases downstream, and this is consistent with influence of stream snowmelt and glacier ice melt documented by Milner *et al.* (2009). It is theoretically consistent with the natural behavior in large Andean basins in humid regions where the pluvial/snow contributor area becomes relatively larger/smaller. The abovementioned suggests that gridded datasets (temperature and rainfall) are an adequate alternative to reproduce the normal hydric regime on large mountainous basins.

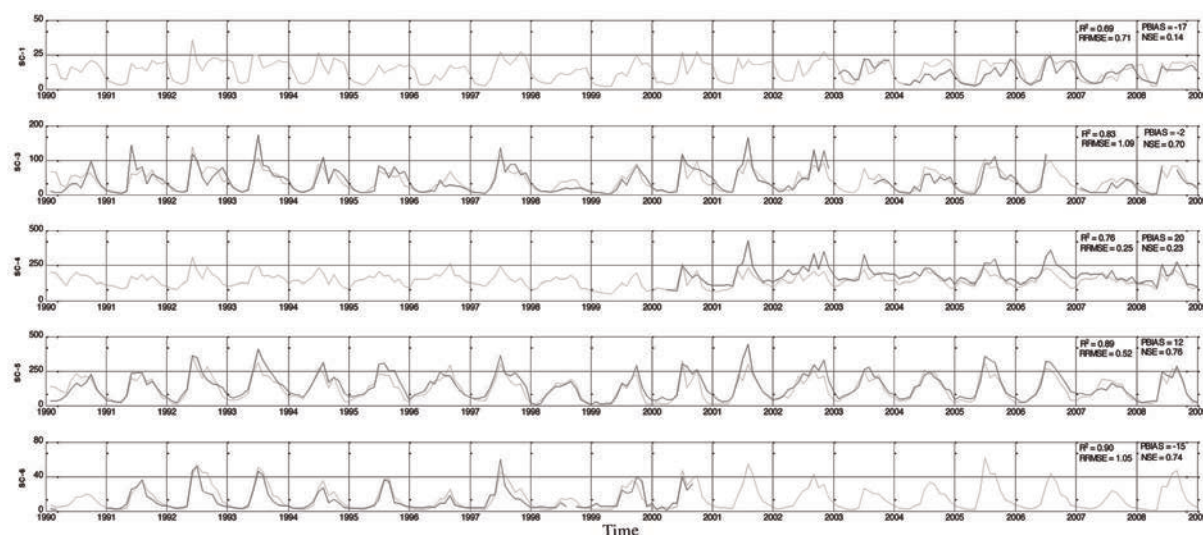
Baseflow and mean flows are well reproduced in both cases, but it is observed that the peak flows in most of the sub-basins during the rainy season are clearly underestimated in the model with UD inputs. This problem is confirmed by the cumulative distribution function plots shown in Figure 6. It can be observed that peaks are underestimated in the SC-1 (as compared to DGA model), SC-3, SC-4, and SC-5 sub-basins. The exception is in the sub-basin located in the Central Valley (SC-6) where peaks are similar, suggesting that UD rainfall data are more representative in smoothed areas. Therefore, UD gridded rainfall dataset deficiency (peak underestimation) is mainly due to the fact that these datasets are interpolated based on at least the 20 rain gauges nearest to the center of the quadrant, producing a smoothed effect; moreover, it is because most of the rainfall stations used in data interpolation are located in areas with low elevation where orographic enhancement is not adequately represented (Nijssen *et al.*, 2001).

Similar results were found by Huffman *et al.* (2001) and Haylock *et al.* (2008) on extreme rainfall gridded datasets using different approaches. On the one hand, Huffman *et al.* (2001) using estimates and an evaluation of gridded daily rainfall data at 1-degree resolution from multisatellite observations mentioned that these appear to be systematically low in complex terrain regions, but they are apparently unbiased in regions of gentler topographic relief. On the other hand, Haylock *et al.* (2008), by comparing observed and gridded extreme daily rainfall and temperature data estimated from land observations, found that spatial interpolation and data resolution (grid size) has a great impact on the magnitude of the extremes by smoothing the peaks, and suggested an inability of the gridded dataset to adequately reproduce climatic spatial variability in complex geographical areas. Furthermore, although UD and GGC datasets are the most detailed gridded data reviewed, and their resolution is better than



DGA_{input}: An input model made up of rainfall data and gridded temperature data measurements; R²: Pearson's correlation coefficient; RRMSE: Relative root mean squared error; NSE: Nash and Sutcliffe coefficient of efficiency; PBIAS: Percent bias.

Figure 4. Observed (dark) and simulated (light) monthly flows (m³ s⁻¹) for the DGA input model.



UD_{input}: An input model made up of gridded rainfall and temperature data; R²: Pearson's correlation coefficient; RRMSE: Relative root mean squared error; NSE: Nash and Sutcliffe coefficient of efficiency; PBIAS: Percent bias.

Figure 5. Observed (dark) and simulated (light) monthly flows (m³ s⁻¹) for the UD input model.

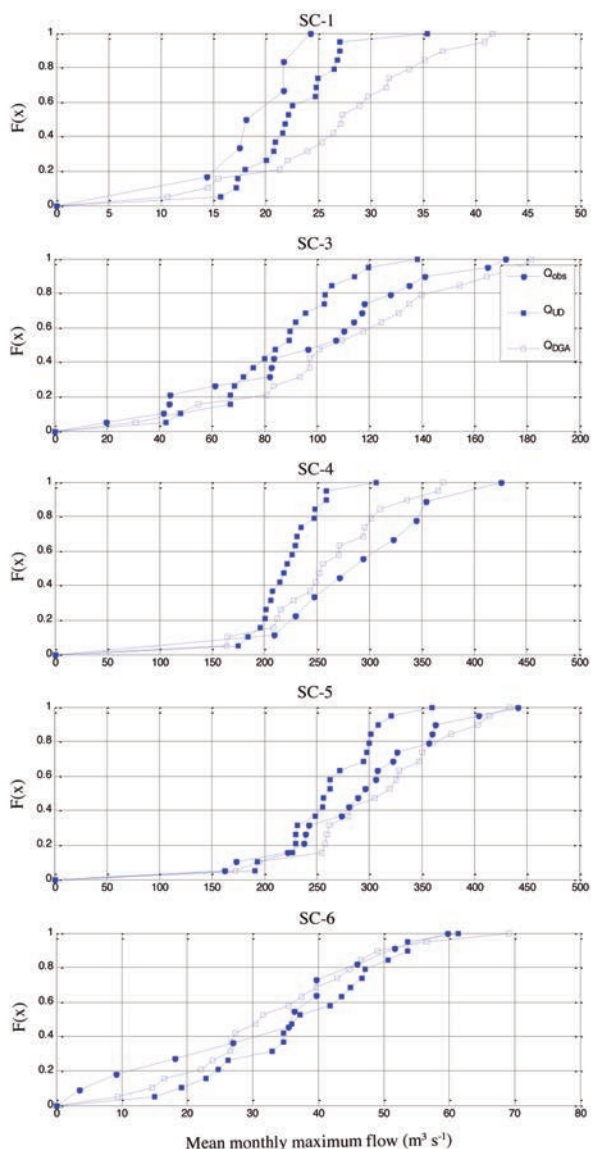
in the abovementioned studies, the problem in complex topographic areas is still observed.

Except for SC-1, which has no lengthy and reliable flow records, all the performance indicators confirm that the model with DGA inputs reproduces the behavior of the Laja basin better than the model with UD inputs, confirming the abovementioned deficiencies in the UD datasets. Moreover, RRMSE allows comparing models and indicates that the UD model has greater mean relative differences than the model built by using measured rainfall data as input.

Following the Motovilov *et al.* (1999) criteria where NSE is categorized as “unsatisfactory” (NSE < 0.36),

“satisfactory” (0.36 < NSE < 0.75), and “good” (NSE > 0.75), SC-4 exhibits a “satisfactory”/“unsatisfactory” model performance with DGA/UD inputs. Meanwhile, following Van Liew *et al.* (2005) where NSE is defined as “unsatisfactory” (PBIAS > |40%|), “satisfactory” (|20%| < PBIAS < |40%|), and “good” (PBIAS < |20%|), SC-4 exhibits a “good” model performance with DGA/UD inputs. Furthermore, if we consider RRMSE and R² values of model behavior, values can be categorized as “good”, which suggests that NSE of SC-4 is affected by its sensitivity to the longitude of the observed data series and to the magnitude of bias (McCuen *et al.*, 2006).

Accordingly, with PBIAS values and criteria by Gupta



Q_{obs} : Observed annual maximum mean monthly flows; Q_{UD} : Annual maximum mean monthly flows simulated by using gridded rainfall and temperature data as input; Q_{DGA} : Annual maximum mean monthly flows simulated by using measured rainfall data and gridded temperature data as input.

Figure 6. Cumulative distribution function plots of the annual maximum mean monthly flows.

et al. (1999), simulated flows show a “good” goodness-of-fit. Moreover, simulated flows are underestimated in SC-4 and SC-5, but overestimated in SC-3. There is also an under/overestimation in SC-6 in the model with DGA/UD input, which is consistent with Figures 4 and 5.

Furthermore, parameter A (parameter to adjust precipitation data) was estimated by ensuring the same amount of precipitation in each sub-basin for both meteorological datasets and during the whole simulation period. As a result, A gave the following values for DGA/UD inputs: i) SC-1 = 1.00/1.45, ii) SC-3 = 1.00/1.47, iii) SC-4 = 1.00/1.22, iv) SC-5 = 1.00/1.01, and v) SC-6 =

0.80/0.87. These values are consistent with orographic effects documented in the Andes Mountains (Falvey and Garreaud, 2007; Favier *et al.*, 2008; Garreaud, 2009) and with the spatially averaged nature of the gridded datasets found. Moreover, UD precipitation datasets are damped and undervalued in terms of the mean, which is mostly noted in the mountainous area (mainly on the Andes slope) where the orographic effect can exert a strong influence on local rainfall patterns (Falvey and Garreaud, 2007).

CONCLUSIONS

Using the UD global monthly gridded climatological datasets as input in a monthly step hydrological model, it is possible to reproduce the behavior of an Andean basin with “good” goodness-of-fit, but with worse results than those datasets using inputs coming directly from rain gauges. However, those differences in the models’ performance are not large, and are mainly due to the fact that precipitation in the UD datasets is clearly underestimated (by 35% on mountainous sub-basins) during the rainy season.

The amounts of UD rainfall datasets in the Andean area of the Laja basin are undervalued and damped, which is observed in the value of parameter A. This effect is mainly due to the large number of stations used in the interpolation, and that most of these stations are located far from the mountains, which results in an inability of this dataset to reproduce the orographic effect, and there is an incidence of underestimation of peak and mean flows, which require input adjustment based on either the data from rain gauges or the results of a long-term mass balance.

In smoothed areas, such as in the Central Valley of south-central Chile, the UD and GGC datasets are useful for performing hydrological studies and reproducing adequately the mean and peak flows, even with similar results than when using measured datasets. Moreover, in complex terrain areas, UD datasets are useful for performing hydrological models where the main objectives are water resource availability studies. Otherwise, for studies relying on peak flows, they are not recommendable due to the damping effect. Thus, if we consider hydrological models related to the main activities of the Laja basin in applications with agricultural purposes, UD inputs can be useful before a treatment. However, for the operation of the Laja hydroelectric plant, UD datasets are useless due to the great underestimation produced by the simulated peak flows.

ACKNOWLEDGEMENTS

This research has been supported by the EU Erasmus Mundus External Cooperation Window, Lote 17, Chile program, by the CONICYT “Programa Nacional de

Formación de Capital Humano Avanzado” scholarship, the BMBF/CONICYT Project 2008099 “Impacto de la variabilidad climática en la disponibilidad de recursos hídricos y requerimientos de riego en la Zona Central de Chile”, and the FONDECYT 11090032 Project “Water and solutes fluxes below irrigated fields: A numerical approach supported by field research to assess the interaction between irrigation and groundwater systems”. The authors thank the Dirección General de Aguas for providing the rain gauge and fluviometric data, and to the anonymous reviewers who helped to improve the paper.

Comparación de datos de precipitación grillados y medidos para estudios hidrológicos a escala de cuenca.

Datos grillados a escala mundial como precipitación y temperatura están siendo cada vez más precisos, accesibles y comunes, pero la utilidad de estos datos y sus limitaciones para estudios hidrológicos, todavía no están bien definidas. En este trabajo se compara el comportamiento de dos modelos hidrológicos, idénticos en estructura, pero construidos con dos entradas diferentes: la precipitación proveniente de estaciones pluviométricas y la precipitación proveniente de datos grillados a escala mundial. El objetivo es evaluar la utilidad de los datos grillados en estudios de disponibilidad de recursos hídricos, principalmente con fines hidroeléctricos y agrícolas. La cuenca andina del río Laja, ubicada en el centro-sur de Chile fue elegida para este estudio. Sobre la base de 18 años de simulación, se concluyó que i) con datos climatológicos grillados en un modelo de balance hídrico mensual es posible reproducir con buena bondad de ajuste el comportamiento de una cuenca andina, pero con peores resultados que usando datos provenientes de estaciones pluviométricas; ii) en términos de montos, los datos grillados en la zona andina de la cuenca del Laja están sub-estimados y amortiguados, efecto que se transfiere a los caudales simulados; y iii) en relación a las principales actividades en la cuenca del Laja, los datos grillados son útiles para aplicaciones hidrológicas con fines agrícolas previo un tratamiento que considere el efecto orográfico. Por otra parte, con fines hidroeléctricos, estos datos no resultan útiles debido a la considerable subestimación de los caudales máximos durante la temporada de lluvia.

Palabras clave: Simulación de caudales, balance hídrico mensual, hidrología superficial, recursos hídricos.

LITERATURE CITED

- CRU. 2000. Climate Research Unit, University of East Anglia, UK. Available at <http://www.cru.uea.ac.uk/cru/data/> (accessed November 2010).
- Falvey, M., and R. Garreaud. 2007. Wintertime precipitation episodes in central Chile: Associated meteorological conditions and orographic influences. *Journal of Hydrometeorology* 8:171-193.
- Favier, V., M. Falvey, A. Rabatel, E. Praderio, and D. López. 2008. Interpreting discrepancies between discharge and precipitation in high-altitude area of Chile's Norte Chico region (26-32° S). *Water Resources Research* 45(2):W02424. doi:10.1029/2008WR006802.
- Fekete, B., C. Vörösmarty, J. Roads, and C. Willmott. 2004. Uncertainties in precipitation and their impacts on runoff estimates. *Journal of Climate* 17:294-304.
- Ferrer, P., E. Brown, y L. Ayala. 1973. Simulación de gastos mensuales en una cuenca pluvial. 12 p. *In* II Coloquio Nacional de la Sociedad Chilena de Ingeniería Hidráulica, Santiago, Chile. Agosto 1973. Sociedad Chilena de Ingeniería Hidráulica, Santiago, Chile.
- Garreaud, R. 2009. The Andes climate and weather. *Advances in Geosciences* 22:3-11.
- GPCC. 2010. Global Precipitation Climatology Centre (GPCC), Offenbach/Main, Germany. Available at http://www.dwd.de/bvbw/appmanager/bvbw/dwdwwwDesktop?_nfpb=true&_pageLabel=_dwdwww_klima_umwelt_klimadaten_deutschland&T82002gsbDocumentPath=Content%2FOeffentlichkeit%2FKU%2FKU2%2FKU21%2FKlimadaten%2Fgerman%2Fteaser2_klimadaten_deutschland.html&_state=maximized&_win dowLabel=T82002&lastPageLabel=_dwdwww_klima_umwelt_klimadaten_deutschland (accessed November 2010).
- GPCP. 2010. Global Precipitation Climatology Project, Greenbelt, USA. Available at <http://precip.gsfc.nasa.gov> (accessed November 2010).
- Gupta, H., S. Sorooshian, and P. Yapo. 1999. Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration. *Journal Hydrologic Engineering* 4(2):135-143.
- Haylock, M., N. Hofstra, A. Klein Tank, E. Klok, P. Jones, and M. New. 2008. A European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006. *Journal of Geophysical Research* 113, D20119. doi:10.1029/2008JD010201.
- Huffman, G., F.R. Adler, M. Morrissey, D. Bolvin, S. Curtis, R. Joyce, et al. 2001. Global precipitation at one-degree daily resolution from multisatellite observations. *Journal of Hydrometeorology* 2:36-50.
- Jiang, T., Y. Chen, X. Chong-yu, X. Chen, X. Chen, and V. Singh. 2007. Comparison of hydrological impacts of climate change simulated by six hydrological models in the Dongjiang Basin, South China. *Journal of Hydrology* 336:316-333.
- Kanamitsu, M., W. Ebisuzaki, J. Woolen, S.-K. Yang, J. Hnilo, M. Fiorino, and G. Potter. 2002. NCEP-DOE AMIP-II Reanalysis (R-2). *Bulletin of the American Meteorological Society* 83:1631-1643.
- Mardones, M., y J. Vargas. 2005. Efectos hidrológicos de los usos eléctricos y agrícolas en la cuenca del Río Laja (Chile Centro-Sur). *Revista de Geografía, Norte Grande* 33:89-102.
- Matsuura, K., and C. Willmott. 2008. Terrestrial air temperature and precipitation: Gridded Monthly and annual time series (1900-2008) Version 2.01. Available at http://climate.geog.udel.edu/~climate/html_pages/download.html#ghcn_T_P2 (accessed November 2010).
- McCuen, R., Z. Knight, and A. Cutter. 2006. Evaluation of the Nash-Sutcliffe efficiency index. *Journal of Hydrologic Engineering* 11:597-602.
- Milner, A., L. Brown, and D. Hannah. 2009. Hydroecological response of river systems to shrinking glaciers. *Hydrological Processes* 23:62-77.
- Motovilov, Y., L. Gottschalk, K. Engeland, and A. Rodhe. 1999. Validation of distributed hydrological model against spatial observations. *Agricultural and Forest Meteorology* 98:257-277.
- Muñoz, E. 2010. Desarrollo de un modelo hidrológico como herramienta de apoyo para la gestión del agua. Aplicación a la cuenca del Río Laja, Chile. 54 p. Tesis de Master. Universidad de Cantabria, Escuela de Caminos Canales y Puertos, Santander, España.
- Nash, J., and J. Sutcliffe. 1970. River flow forecasting through conceptual models part I: A discussion of principles. *Journal of Hydrology* 10(3):282-290.
- NCEP. 2004. The NCEP/DOE AMIP-II Reanalysis (Reanalysis-II). *In* National Centers for Environmental Prediction, Camp Springs, Maryland, USA. Available at <http://dss.ucar.edu/pub/reanalysis2> (accessed November 2010).

- New, M., M. Hume, and P. Jones. 1999. Representing twentieth century space-time climate variability: I. Development of a 1961-1990 mean monthly terrestrial climatology. *Journal of Climatology* 12:829-856.
- New, M., M. Hume, and P. Jones. 2000. Representing twentieth century space-time climate variability: II. Development of 1901-1996 monthly grids of terrestrial surface. *Journal of Climatology* 13:2217-2238.
- Nijssen, B., G. O'Donnell, D. Lettenmaier, D. Lohmann, and E. Wood. 2001. Predicting the discharge of global rivers. *Journal of Climate* 14:3307-3323.
- Oki, T., and S. Kanae. 2006. Global hydrological cycles and world water resources. *Freshwater Resources* 313:1068-1072.
- Pitcher, E., R. Malone, V. Ramanathan, M. Blackmon, K. Puri, and W. Bourke. 1983. January and July simulations with a Spectral General Circulation Model. *Journal of the Atmospheric Sciences* 40:580-604.
- Rijsberman, F. 2005. Water scarcity: Fact or fiction? *Agricultural Water Management* 80:5-22.
- Stehr, A., P. Debels, F. Romero, and H. Alcayaga. 2008. Hydrological modeling with SWAT under conditions of limited data availability: Evaluation of results from a Chilean case study. *Hydrological Sciences* 53(3):588-601.
- Thiele, R., H. Moreno, S. Elgueta, A. Lahsen, S. Rebolledo, y M. Petit-Breuihl. 1998. Evolución geológico-geomorfológica cuaternaria del tramo superior del valle del río Laja. *Revista Geológica de Chile* 25:229-253.
- Thornthwaite, C. 1948. An approach toward a rational classification of climate. *Geographical Review* 38:55-94.
- Van Liew, M., J. Arnold, and D. Bosch. 2005. Problems and potential of autocalibrating a hydrologic model. *Transaction of the American Society of Agricultural Engineering* 48:1025-1040.
- Vicuña, S., R. Garreaud, and J. McPhee. 2011. Climate Change Impacts on the hydrology of a snowmelt driven basin in semiarid Chile. *Climatic Change* 105:469-488.
- Willmott, C., and D. Legates. 1991. Rising estimates of terrestrial and global precipitation. *Climate Research* 1:179-186.